JOURNAL OF APPROXIMATION THEORY **23**, 73–77 (1978)

A Note on the Method of Steepest Descent

ZITA DIVIŠ

Department of Mathematics, The Ohio State University, Columbus, Ohio 43210

Communicated by R. Bojanic

Received October 1, 1976

Consider a differentiable functional (in the sense of Gateaux

$$f: B \to R$$

where B is a real Banach space. Assume that f attains its minimum at a unique point $x_0 \in B$, so that

$$F(x_0) = \inf_{x \in B} F(x) > -\infty.$$

In this paper we are interested in approximations to x_0. There are few instances where sequences converging to x_0 can be actually constructed. We shall investigate the convergence of approximation to x_0 obtained by the method of steepest descent.

Throughout the paper we shall assume that B is a real reflexive Banach space and the pairing between $y \in B^*$ and $x \in B$ will be denoted by (y, x). Further, we shall denote $F = \operatorname{grad} f$ and assume that this operator $F: B \to B^*$ is monotone, i.e.,

$$(F(x) - F(y), x - y) \geqslant 0$$

for each $x, y \in B$.

The sequence of successive approximations to x_0 is defined as follows:

$$x_{n+1} = x_n - \epsilon_n AF(x_n), \qquad n = 1, 2, \ldots, \tag{1}$$

where x_1, $A: B^* \to B$ and the "relaxation" coefficients ϵ_n will be specified later. Such processes were studied by many authors. Our note concerns the results of Vajnberg [4, 5]. For a large bibiliography on the subject the reader may consult the article by Ljubič and Majstrovskij [3].

Before we discuss the convergence of the process (1) we shall state two auxiliary results guaranteeing existence and uniqueness of $\min_{x \in B} f(x)$.

PROPOSITION 1. *Assume*:

(i) *F is strictly monotone, i.e.,* $(F(x) - F(y), x - y) > 0$ *for all* $x, y \in B$.

0021-9045/78/0231–0073$02.00/0

(ii) *Let* λ *be a real-valued, measurable function defined on* $[0, +\infty)$ *such that* $(F(x), x) \geqslant \lambda(\| x \|)$.

(iii) *For certain* $R_0 > 0$, $0 < \int_0^{R_0} | \lambda(t)|/t \, dt < +\infty$.

Then there exists a unique point $x_0 \in B$ *such that* $f(x_0) = \inf_{x \in B} f(x) > -\infty$.

For the proof see Diviš [2].

In particular, we have the following

PROPOSITION 2. *Let* (i), (ii) *of the above proposition be satisfied. Instead of* (iii) *we shall assume that* $\lambda(t)/t$ *is integrable on* $(0, R)$ *for every* $R > 0$ *and*

$$\lim_{R \to +\infty} \int_0^R \frac{\lambda(t)}{t} \, dt = +\infty. \tag{2}$$

Then $\inf_{x \in B} f(x)$ *is assumed at a unique point* $x_0 \in B$ *and, moreover,* $\lim_{\|x\| \to +\infty} f(x) = +\infty$.

We shall prove this last assertion. We have $f(x) = f(0) + \int_0^1 (F(tx), tx) \times (dt/t) \geqslant f(0) + \int_0^1 (\lambda(t \cdot \| x \|)/t) \, dt$ and for $\| x \| = R$ we obtain $f(x) \geqslant f(0) + \int_0^R (\lambda(t)/t) \, dt$, whence $\lim_{\|x\| \to +\infty} f(x) = +\infty$ by (2).

Remark 1. Let φ: $[0, +\infty) \to [0, +\infty)$ be a strictly increasing continuous function, $\varphi(0) = 0$, $\varphi(r) \to +\infty$ as $r \to +\infty$. Let B be a strictly convex space. Then A: $B^* \to B$ is a duality mapping corresponding to φ if

$$(y, Ay) = \| y \| \cdot \| Ay \|, \qquad \| Ay \| = \varphi(\| y \|).$$

For duality mappings, see e.g., Browder [1].

Our results can now be formulated as follows.

THEOREM 1. *Assume*:

(i) *F satisfies all the hypotheses of Proposition* 1 *and further let F be* δ*-Hölder continuous with some* $0 < \delta \leqslant 1$, *i.e.*,

$$\| F(x + h) - F(x) \| \leqslant M(r) \cdot \| h \|^{\delta}, \tag{3}$$

where M is nondecreasing positive continuous function defined on $[0, +\infty)$; $x, x + h \in D_r = \{x: \| x \| \leqslant r\}$. *Moreover, let*

$$(F(x + h) - F(x), h) \geqslant \mathfrak{S}(\| h \|) \tag{4}$$

for every $x, h \in B$, *where* $\mathfrak{S}$ *is a continuous strictly increasing function defined on* $[0, +\infty)$ *and* $\mathfrak{S}(0) = 0$.

(ii) *B is strictly convex and $A: B^* \to B$ be a duality corresponding to a function φ described in Remark* 1 *for which*

$$\overline{\lim_{r\to 0+}}\ \varphi^{\delta}(r)/r < +\infty. \tag{5}$$

(iii) *Let ϵ_n be such that*

$$\tfrac{1}{4} \leqslant \epsilon_n{}^{\delta} M_n P_n \leqslant \tfrac{1}{2}, \tag{6}$$

where $M_n = \max(1, M(R_n))$, $P_n = \max(1, \varphi^{\delta}(\| F(x_n)\|)/\| F(x_n)\|)$, *and* $R_n \geqslant \| x_n \| + \varphi(\| F(x_n)\|)$. *Then the sequence* (1) *converges towards* x_0 *with the choice* $x_1 = 0$.

THEOREM 2. *Assume that F satisfies all the hypotheses of Proposition* 2, *the inequalities* (3), (4) *and the hypotheses* (ii), (iii) *of Theorem* 1. *Then the process* (1) *converges to x_0 with x_1 chosen arbitrarily.*

Remark 2. The result of Vajnberg [5] is contained in Theorem 2 with the special choice of $\varphi(r) = r$, $\delta = 1$ and $\mathfrak{S}(r) = r.\ \mathfrak{S}_0(r)$, $r \geqslant 0$, where $\mathfrak{S}_0(r)$ satisfies such conditions as we impose on $\mathfrak{S}$. The proof will be a modification of the method used by Vajnberg in [5].

Proof of Theorem 1. By Proposition 1, f assumes a minimum in E at a unique point x_0. We have then $F(x_0) = 0$. Further, we have the estimate

$$f(x) > f(0) + \int_0^{R_0} \frac{\lambda(t)}{t}\, dt$$

for all $x \in \{x: \| x \| = R_0\}$. It follows then that

$$f(x) > f(0)$$

for all $x \in E$, $\| x \| \geqslant R_0$. (For proof see Diviš [2].) Now let $x_1 = 0$ and consider the difference $f(x_n) - f(x_{n+1})$. Using Lagrange's formula, there exists a $\tau_n \in (0, 1)$ such that $f(x_n) - f(x_{n+1}) = (F(x_{n+1} + \tau_n(x_n - x_{n+1})), x_n - x_{n+1})$. Using (3) and the definition of A we obtain

$$\begin{aligned}
&f(x_n) - f(x_{n+1})\\
&\quad = (F(x_n), x_n - x_{n+1}) - (F(x_{n+1} + \tau_n(x_n - x_{n+1})) - F(x_n), x_{n+1} - x_n)\\
&\quad = \epsilon_n(F(x_n), AF(x_n)) - (F(x_{n+1} + \tau_n(x_n - x_{n+1})) - F(x_n), x_{n+1} - x_n)\\
&\quad \geqslant \epsilon_n \| F(x_n)\| \cdot \varphi(\| F(x_n)\|)\\
&\qquad - \| F(x_{n+1} + \tau_n(x_n - x_{n+1})) - F(x_n)\| \cdot \| x_{n+1} - x_n \|\\
&\quad \geqslant \epsilon_n \| F(x_n)\| \cdot \varphi(\| F(x_n)\|) - M(R_n) \cdot (1 - \tau_n)^{\delta} \cdot \| x_{n+1} - x_n \|^{1+\delta}\\
&\quad \geqslant \epsilon_n \| F(x_n)\|\, \varphi(\| F(x_n)\|) - M(R_n) \cdot \epsilon_n^{1+\delta} \cdot \varphi^{1+\delta}(\| F(x_n)\|).
\end{aligned}$$

Notice that indeed x_n, $x_{n+1} + \tau_n(x_n - x_{n+1}) \in D_{R_n}$ for $n = 1, 2, \ldots$. We have namely $\| x_n \| \leqslant R_n$, $\| x_{n+1} \| \leqslant \| x_n \| + \epsilon_n \varphi(\| F(x_n)\|)$ and by (6), $\epsilon_n \leqslant 1$. Here we assume without loss of generality that $F(x_n) \neq 0$ for every $n = 1, 2, \ldots$, We shall write now $M_n = \max(1, M(R_n))$ and $P_n = \max(1, \varphi^\delta(\| F(x_n)\|)/\| F(x_n)\|)$. Observe that according to the Remark 1 and inequality (5), if $\{\| F(x_n)\|\}$ is a bounded sequence then so is $\{P_n\}$.

With the above notation we have $f(x_n) - f(x_{n+1}) \geqslant \epsilon_n \| F(x_n)\| \cdot \varphi(\| F(x_n)\|)\ [1 - M_n \epsilon_n{}^\delta P_n]$. Consequently, if $\epsilon_n{}^\delta$ satisfy the inequalities (6), then

$$f(x_n) - f(x_{n+1}) \geqslant \tfrac{1}{2}\epsilon_n \| F(x_n)\| \cdot \varphi(\| F(x_n)\|) > 0.$$

Hence, the sequence $\{ f(x_n)\}$ is decreasing and $\lim_{n\to\infty} f(x_n) \geqslant \inf_{x\in B} f(x)$. With the choice $x_1 = 0$ we find

$$f(x_1) = f(0) > f(x_2) > f(x_3) > \cdots.$$

But for all x, $\| x \| \geqslant R_0$, we have $f(x) > f(0)$ as mentioned above. Thus, all the x_n's must lie in D_{R_0} and we conclude $\| x_n \| \leqslant R_0$ $(n = 1, 2, \ldots)$. Consequently, by (3) $\{\| F(x_n)\|\}$ is bounded. Thus, both $\{M_n\}$ and $\{P_n\}$ are bounded sequences and let $1 \leqslant M_n \leqslant K$, $1 \leqslant P_n \leqslant K$. Then

$$1/4M_nP_n \geqslant 1/4K^2 > 0$$

and taking into account that $\lim_{n\to\infty} (f(x_n) - f(x_{n+1})) = 0$ and that

$$\begin{aligned} f(x_n) - f(x_{n+1}) &\geqslant \tfrac{1}{2}\epsilon_n \| F(x_n)\|\ \varphi(\| F(x_n)\|) \\ &\geqslant \tfrac{1}{2} \cdot (1/4K^2)^{1/\delta} \cdot \| F(x_n)\| \cdot \varphi(\| F(x_n)\|) \end{aligned}$$

we conclude that

$$\lim_{n\to\infty} F(x_n) = 0.$$

Next, using (4) and the fact that $F(x_0) = 0$, we estimate

$$\begin{aligned} \mathfrak{S}(\| x_n - x_0 \| &\leqslant (F(x_n) - F(x_0), x_n - x_0) \\ &\leqslant \| F(x_n)\| \cdot \| x_n - x_0 \| \leqslant 2R_0 \cdot \| F(x_n)\| \end{aligned}$$

whence $\mathfrak{S}(\| x_n - x_0 \|) \to 0$ as $n \to +\infty$ and then, $\| x_n - x_0 \| \to 0$ as $n \to +\infty$. We have, in fact, the error estimate

$$\| x_n - x_0 \| \leqslant \mathfrak{S}^{-1}(2R_0 \cdot \| F(x_n)\|)$$

and also

$$f(x_n) - f(x_0) = (F(x_0 + \tau_n(x_n - x_0)), x_n - x_0) \leqslant C \cdot \| x_n - x_0 \|^{1+\delta},$$

C independent of n.

Remark 3. From the last inequality we see, that $\{x_n\}$, under the circumstances, is a minimizing sequence for f.

Proof of Theorem 2. It differs from the above proof in the following way: First, $x_1 \in B$ can be chosen arbitrarily. Next, the boundedness of the sequence $\{x_n\}$ must be shown in a different way. We have, as before,

$$f(x_n) > f(x_{n+1}) \geqslant \inf_{x \in B} f(x) > -\infty, \qquad n = 1, 2, \ldots;$$

thus $\lim_{n \to \infty} f(x_n)$ exists and is finite. If $\{x_n\}$ was not bounded, for a subsequence $\{x_{n_k}\}$ of $\{x_n\}$, $\| x_{n_k} \| \to +\infty$ and we would have $\lim_{k \to +\infty} f(x_{n_k}) = +\infty$, a contradiction. This finishes the proof. Note that for the error estimate we again obtain

$$\| x_n - x_0 \| \leqslant \mathfrak{S}^{-1}(2C \cdot \| F(x_n) \|),$$

where $\| x_n \| \leqslant C$ $(n = 1, 2, \ldots)$.

References

1. F. E. Browder, Problèmes non linéaires; Lecture Notes, Université de Montréal, 1966.
2. Z. Diviš, Existence theorem and convergence of minimizing sequences in extremum problems, *SIAM J. Math. Anal.* **2** (1971), 505–510.
3. J. L. Ljubič and G. D. Majstrovskij, General theory of relaxation processes for convex functionals, *Usp. Mat. Nauk* **25** (1970), 57–113.
4. M. M. Vajnberg, On the convergence of the process of the steepest descent for nonlinear equations, *Sibirsk. Mat. Ž.* **2** (1961), 201–221.
5. M. M. Vajnberg, Variational Method and the Method of Monotone Operators (Russian), pp. 154–161, Moscow, 1972.